

RWLP0077

# **Application of Biomarkers in Cancer Epidemiology**

Edited by P. Toniolo, P. Boffetta, D.E.G. Shuker,  
N. Rothman, B. Hulka and N. Pearce

IARC Scientific Publications No. 142

International Agency for Research on Cancer, Lyon, 1997

PM3006483655

# General issues of study design and analysis in the use of biomarkers in cancer epidemiology

N. Pearce and P. Boffetta

Other contributions to this volume have discussed sources of variation (see Vineis) and measurement error (see White). In this article, we focus on statistical issues involved in the design and analysis of epidemiological studies that use biomarkers. We do not consider statistical issues of laboratory analyses.

In general, epidemiological research involves studying the external, modifiable causes of diseases in populations (McMichael, 1994, 1995) with the intention of developing preventive interventions. In some instances, this activity can be enhanced by using internal biomarkers to obtain better measurements of internal exposure (dose), disease or individual susceptibility. Statistical issues in studies using biomarkers of exposure are not markedly different from those involved in other epidemiological studies based on measures of external exposure, while studies using biomarkers of disease pose specific problems due to the lack of persistence of some of these markers, and the analysis of interaction is of particular interest in studies using markers of susceptibility. As with other epidemiological studies, the statistical analysis of a study involving biomarkers involves in general: (1) relating a particular disease (or health outcome, such as a marker of early effect) to (2) a particular exposure while (3) controlling for systematic error, (4) assessing interactions with other exposures and (5) assessing the possibility of random error. We will consider each of these five aspects of study design and analysis in turn. We will restrict our discussion to full-scale epidemiological studies, i.e. we assume that the use of biomarkers, be they of exposure, effect or susceptibility, aims to contribute to the elucidation of the causal relationships in human populations between diseases and factors such as external exposures, personal habits, genetic traits and interventions. Issues in the design and analysis of transitional studies, in which the main aim is the validation of the markers themselves, are outside the scope of this chapter.

## Measuring disease with biomarkers

Epidemiological studies are usually based on a particular population followed over a particular period of time. Miettinen (1985) has termed this study population the 'base population' and its experience over time the 'study base'. The different epidemiological study designs differ only in the manner in which the study base is defined and the manner in which information is drawn from the study base (Checkoway *et al.*, 1989). Thus, epidemiological studies may involve measuring either incidence or prevalence of disease. This distinction is important when a biomarker is being used to measure either the disease under study or early biological effects that are considered to be valid predictors of disease risk (e.g. Rothman *et al.*, 1995). In particular, many studies measuring disease with biomarkers are of cross-sectional design and measure the prevalence of the disease, which is dependent on its incidence and its duration. Thus, in a study looking at markers of cell damage as an effect of exposure to known or suspected carcinogens, the results would depend on factors such as the turnover of the cells in which the marker is measured or the capacity to repair the damage. For example, there is evidence that chromosomal aberrations caused by some carcinogens, such as arsenic and benzene, last for longer periods than aberrations caused by vinyl chloride, and the reason for this difference is not known (Schwartz, 1990).

## Incidence studies

Three measures of disease incidence are commonly used in incidence studies. The (person-time) incidence rate (or incidence density; Miettinen,

1985) is a measure of the disease occurrence per unit time. A second measure of disease occurrence is the cumulative incidence (incidence proportion; Miettinen, 1985) or risk, which is the proportion of study subjects who experience the outcome of interest at any time during the follow-up period. A third possible measure of disease occurrence is the incidence odds (Greenland, 1987), which is the ratio of the number of subjects who experience the outcome to the number who do not experience the outcome. As for the cumulative incidence, the incidence odds is dimensionless, but it is necessary to specify the time period over which it is being measured. In incidence studies involving biomarkers of disease, it is therefore important to consider whether a particular biomarker measures incidence or cumulative incidence.

Corresponding to these three measures of disease occurrence, there are three principal ratio measures of effect that can be used in cohort studies. The measure of primary interest is often the rate ratio (incidence density ratio), which is the ratio of the incidence rate in the exposed group to that in the non-exposed group. A second commonly used effect measure is the risk ratio (cumulative incidence ratio), which is the ratio of the cumulative incidence in the exposed group to that in the non-exposed group. When the outcome is rare over the follow-up period, the risk ratio is approximately equal to the rate ratio. A third possible effect measure is the incidence odds ratio, which is the ratio of the incidence odds in the exposed group to that in the non-exposed group. An analogous approach can be used to calculate measures of effect based on the differences rather than the ratios, in particular the rate difference and the risk difference.

In incidence studies involving biomarkers of disease, it is therefore important to consider whether a particular biomarker measures incidence or cumulative incidence. For example, if the 'disease' under study is hepatitis B virus (HBV) infection, then a survey of the prevalence of HBV markers, in a cohort that has been followed over time, will indicate the cumulative incidence of infection in the cohort (with the exception of those who have died from any cause during follow-up or who no longer show evidence of infection). It will not directly indicate the incidence rate of infections; this would require repeated prevalence surveys over time.

#### Incidence case-control studies

Incidence case-control studies involve studying all of the incident cases of disease generated by the study base and a control group sampled at random from the same study base. The relative risk measure is the incidence odds ratio; the effect measure that this estimates depends on the manner in which controls are selected. Once again, there are three main options (Pearce, 1993).

One option is to select controls from those who do not experience the outcome during the follow-up period, i.e. the survivors (those who did not develop the outcome at any time during the follow-up period). In this instance, a sample of controls chosen by cumulative incidence sampling will estimate the exposure odds of the survivors, and the odds ratio obtained in the case-control study will therefore estimate the incidence odds ratio in the base population. Controls can also be sampled from the entire base population (those at risk at the beginning of follow-up), rather than just from the survivors (those at risk at the end of follow-up). In such case-base sampling, the controls will estimate the exposure odds in the base population of persons at risk at the start of follow-up, and the odds ratio obtained in the case-control study will therefore estimate the risk ratio in the base population. The third approach is to select controls longitudinally throughout the course of the study (Miettinen, 1976); this is sometimes described as 'risk-set sampling' (Robins *et al.*, 1986), 'sampling from the study base' (the person-time experience; Miettinen, 1985) or 'density sampling' (Kleinbaum *et al.*, 1982). In this instance, the controls will estimate the exposure odds in the study base, and the odds ratio obtained in the case-control study will therefore estimate the rate ratio in the study base.

In incidence case-control studies involving biomarkers of disease, it is therefore important to consider whether a particular biomarker measures incidence or cumulative incidence. These issues determine not only which measure of effect is being estimated, but also which method of control selection is appropriate, and the resulting methods of data analysis.

#### Prevalence studies

The term prevalence denotes the number of cases of disease existing in the population at the time the study was conducted. If we denote the prevalence

of a disease in the study population as  $P$ , assume that the incidence rate is  $I$ , the population has size  $N$ , and there is no migration in or out of the prevalence pool, then it can be shown that the prevalence odds ratio ( $P$ ) multiplied by the incidence odds ratio ( $D$ ):

$$P/(1 - P) = I \times D$$

Thus, the prevalence odds ratio is proportional to the disease incidence odds ratio is estimated to

$$OR = I_1 D_1 / I_0 D_0$$

An increased prevalence of disease reflects the influence of factors affecting disease incidence and/or factors affecting disease duration. The different measures of disease incidence and prevalence are likely to involve different etiologies and disease (see below for a discussion of the likely need to require different methods of the etiologically important factors).

#### Prevalence case-control studies

Just as an incidence case-control study can be used to obtain the same findings as a prevalence case-control study, a prevalence case-control study can be used to obtain the same findings as an incidence case-control study in a more efficient manner. In a prevalence case-control study, exposure information is obtained from serum samples, and the controls are selected by obtaining exposure information from prevalent cases of the disease. In this instance, the controls will estimate the exposure odds in the study base, and the odds ratio obtained in the case-control study will therefore estimate the prevalence odds ratio in the study base. In turn, the prevalence odds ratio estimates the incidence odds ratio that the average duration of disease in the exposed and non-exposed groups is the same. An increased prevalence of disease reflects the influence of factors that increase the incidence and/or factors that increase the duration of disease.

### studies

studies involve studying all of disease generated by the of group sampled at random se. The relative risk measure tio; the effect measure that s on the manner in which ynce again, there are three 1993).

ct controls from those who outcome during the follow- vitors (those who did not ny time during the follow- nce, a sample of controls incidence sampling will es- ls of the survivors, and the he case-control study will incidence odds ratio in the ols can also be sampled pulation (those at risk at up), rather than just from k at the end of follow-up). ng, the controls will esti- n the base population of rt of follow-up, and the e case-control study will k ratio in the base popu- ch is to select controls t the course of the study s sometimes described as s *et al.*, 1986), 'sampling person-time experience; y sampling' (Kleinbaum ce, the controls will esti- the study base, and the case-control study will ratio in the study base. l studies involving bio- efore important to con- r biomarker measures nce. These issues h measure of effect is hich method of control l the resulting methods

is the number of cases pulation at the time the denote the prevalence

of a disease in the study population by  $P$ , and if we assume that the incidence rate is constant over time, the population has reached a 'steady state' and there is no migration into or out of the prevalence pool, then it can be shown (Rothman, 1986) that the prevalence odds is equal to the incidence rate ( $I$ ) multiplied by the average disease duration ( $D$ ):

$$P/(1 - P) = I \times D$$

Thus, the prevalence odds is directly proportional to the disease incidence, and the prevalence odds ratio is estimated to be:

$$OR = I_1 D_1 / I_0 D_0$$

An increased prevalence odds ratio may thus reflect the influence of factors that increase the disease incidence and/or factors that increase disease duration. The different mechanisms involved in increasing disease incidence or disease duration are likely to involve different time patterns of exposure and disease (see below), which in turn are likely to require different biomarkers for measurement of the etiologically relevant exposures.

### Prevalence case-control studies

Just as an incidence case-control study can be used to obtain the same findings as a full cohort study, a prevalence case-control study can be used to obtain the same findings as a full prevalence study in a more efficient manner. In particular, if obtaining exposure information is difficult or costly (e.g. if it involves serum samples), then it may be more efficient to conduct a prevalence case-control study by obtaining exposure information on all of the prevalent cases of the disease under study and a sample of controls selected at random from the non-cases. In this instance, a sample of controls will estimate the exposure odds of the non-cases, and the odds ratio obtained in the prevalence case-control study will therefore estimate the prevalence odds ratio in the base population, which in turn estimates the incidence rate ratio, provided that the average duration of disease is the same in the exposed and non-exposed groups. Once again, an increased prevalence odds ratio may reflect the influence of factors that increase disease incidence and/or factors that increase disease duration, and

the different mechanisms involved are likely to require different biomarkers for measurement of the etiologically relevant exposures.

### Measuring exposure with biomarkers

#### Validity of biomarkers of exposure

There are considerable shortcomings in many currently available biomarkers of exposure, including problems of measuring historical exposures; uncertainties as to what a biomarker is measuring; greater susceptibility to confounding in some instances; problems of application to public health policy (Pearce *et al.*, 1995); the disease process affecting the level of the biomarker; and problems of validity of laboratory measurements (Boffetta, 1995). These issues are covered elsewhere in this volume. In this section we concentrate on issues of study design and analysis when measuring exposure with biomarkers, particularly with regard to time-related exposures with a relatively long induction and latency period between exposure and the subsequent occurrence of disease (as is the situation in most cancer epidemiology studies). The issues we discuss are not unique to a particular study design (cohort studies, case-control studies, cross-sectional studies) but rather apply to all studies in which the etiologically relevant time period involves a relatively long induction and latency period, thereby posing problems with the measurement of exposure during this period.

#### Time-related exposures

Some biomarkers measure factors that are fixed and do not change over time in an individual, e.g. genetic susceptibility genes that may interact with xenobiotic factors in cancer causation (Rothman, 1995). Other biomarkers measure factors that change over time, e.g. micronutrient levels in serum may change from day-to-day (Willett, 1990).

In studies (both prospective and retrospective) of long-term health effects involving time-related exposures, it is important that the time patterns of the study exposure and of the relevant confounders should be taken into account in the analysis (Pearce *et al.*, 1986). In particular, it is important that the principal exposure under study should be analysed in a time-related manner, taking account of the likely induction and latency periods, and the relative etiological importance of exposure intensity, exposure duration and

cumulative exposure. The simplest approach is to analyse the cumulative exposure in a time-related manner, and this may suffice when the aim is merely to consider whether or not there is an effect of exposure. However, once it has been provisionally assumed that an effect exists, attention then shifts to understanding the nature of the effect. In this context, the temporal pattern of exposure and outcome can be considered by examining the effects of exposures in specific time windows while controlling for time-related confounders and for the effects of exposures in other time windows. A more sophisticated approach is direct fitting of a theoretical model of carcinogenesis (Pearce, 1992), which requires assumptions as to the relevance of the times in which the exposure occurred and in which the marker was measured. This may be particularly relevant in studies including the measurement of the exposure, the marker and the disease, which are mainly aimed at elucidating the role of the marker in the exposure-disease relationship (Schatzkin *et al.*, 1993). An example of this type of study is the investigation of the role of human papilloma virus (HPV) in the etiology of cervical cancer (the 'exposure' in this case being factors such as number of sexual partners and age at first intercourse). In this case, the association with HPV infection, when properly measured with PCR-based assays, is of such a magnitude that there is little concern about the relevance of the time periods (Muñoz *et al.*, 1992). However, for cancers from other organs, the association with HPV infection is less clear-cut, and the relevance of the timing of the infection may be one of the unknown factors modifying the exposure-marker-disease relationship.

Thus, biological measures of time-related exposures must be able to measure changes in exposure levels over time. In particular, stored biological samples may not provide valid measurements of long-term patterns of exposure when there are significant variations in exposure over time, unless samples have been taken repeatedly over the course of the study (Armstrong *et al.*, 1992). If it is not possible to take repeated biological samples over time, then it is essential that the samples that are taken relate to the etiologically relevant time period.

Many currently available biomarkers only indicate relatively recent exposures. For example, it is well known that serum levels of micronutrients re-

flect recent rather than historical dietary intake (Willett, 1990); given the long induction time of most cancers it is usually exposures between 10 and 30 years previously that are etiologically relevant. While this may not be a limitation in cross-sectional studies, provided that the etiologically relevant time period is close to the time of data collection, it is an important limitation in cohort and case-control studies aiming to assess the effects of historical exposures. Some biomarkers are better than others in this respect (particularly biomarkers for exposure to biological agents), but even the best markers of chemical exposures reflect only the last few weeks or months of exposure. On the other hand, with some biomarkers (e.g. serum levels of TCDD; Johnson *et al.*, 1992) it may be possible to estimate historical levels if the exposure period is known, if the half-life is relatively long (and is known) and if it is assumed that no significant exposure has occurred more recently, or if it is reasonable to assume that exposure levels have remained stable over time.

However, historical information on exposure surrogates will often be more valid than current direct measurements of exposure or dose. This situation has long been recognized in occupational epidemiology, where the use of work history records in combination with a job-exposure matrix (based on historical exposure measurements of work areas rather than individuals) is usually more valid than current exposure measurements (whether based on environmental measurements or biomarkers) because of changes in exposure levels over time (Checkoway *et al.*, 1989). Similar problems may occur in the measurement of other carcinogenic exposures. For example, even the best currently available measures of exposure to tobacco smoke, such as plasma or urinary cotinine, appear to have similar validity to questionnaires for the measurement of current exposures; their very short half-life makes them inferior to questionnaires in the estimation of historical exposures (Pearce *et al.*, 1995). On the other hand, some biomarkers would appear to have value in the validation of questionnaires (Forastiere *et al.*, 1993), which can then be used to estimate historical exposures.

Another example in which timing of sample collection is of great importance is in the case of DNA adducts (Wilcosky & Griffith, 1990). Since most adducts are readily repaired, any measure of

exposure based on DN the time between the collection; this pattern factors such as the ac which in turn may also fluence on the outcome with case-control status and repair are particular

- the extent to which measured represents active adducts formed usually not known; a
- in the case of measurement (or immediate end of a working day much of the adducts enough to be biological

**Analysis based on pooled samples.** These issues of the timing of particular concern in samples. If the samples same etiologically relevant pooled sample will represent samples taken from a variety there may be considerable regard to the exposure relevant time period.

#### Systematic error

The major possible types are the same in traditional studies involving biomarkers various types of bias can major classes: selection bias confounding (Rothman, intended to give a complete types of bias; rather, we involving data analysis, small sets of samples stratified since collection; however, substantially reduce the advantage

#### Selection bias

Selection bias involves biases by which the study from the study base. Selected by including all of the subjects (study) and obtaining a representative

historical dietary intake long induction time of exposures between 10 years are etiologically relevant. It may be a limitation in cross-sectional studies that the etiologically relevant time to the time of data collection is long. This is a limitation in cohort and case-control studies to assess the effects of exposures. Biomarkers are better than clinical endpoints (e.g. serum levels of enzymes, antibodies, etc.) particularly biomarkers (e.g. serum levels of enzymes, antibodies, etc.) (e.g. serum levels of enzymes, antibodies, etc.) it may be possible to measure the exposure period is relatively long (and is not known) and is not known that no significant difference recently, or if it is not known exposure levels have re-

information on exposure are more valid than current exposure or dose. This is not true in occupational use of work history or a job-exposure matrix. Measurements of work history is usually more valid than measurements (whether based on self-reports or biomarkers) of exposure levels over time. Similar problems may exist for other carcinogenic exposures. The best currently available data on tobacco smoke, asbestos, and radon, appear to have been derived from the measurement of their very short half-lives in the environment (Pearce *et al.*, 1990). Biomarkers would be useful in the validation of questionnaires which can then be used to measure exposure.

The timing of sample collection is in the case of cross-sectional studies (Hoff, 1990). Since the timing of sample collection is in the case of cross-sectional studies, any measure of

exposure based on DNA adducts will depend on the time between the end of exposure and sample collection; this pattern can then be modified by factors such as the activity of repair enzymes, which in turn may also have an independent influence on the outcome, i.e. may be associated with case-control status. DNA adduct formation and repair are particularly problematic, since:

- the extent to which the amount of adducts measured represents the amount of biologically active adducts formed (as discussed above) is usually not known; and
- in the case of measurements taken during exposure (or immediately thereafter, such as at the end of a working day) it is not known how much of the adducts found would persist long enough to be biologically important.

#### *Analysis based on pooled samples*

These issues of the timing of sample collection are of particular concern in analyses based on pooled samples. If the samples were not all taken at the same etiologically relevant time period, then the pooled sample will represent the average of samples taken from a variety of time periods. Thus, there may be considerable misclassification with regard to the exposure levels at the etiologically relevant time period.

#### **Systematic error**

The major possible types of systematic error (bias) are the same in traditional epidemiology and in studies involving biomarkers (Boffetta, 1995). The various types of bias can be grouped into three major classes: selection bias, information bias, and confounding (Rothman, 1986). This section is not intended to give a comprehensive review of these types of bias; rather, we will concentrate on issues involving data analysis. One solution is to pool small sets of samples stratified on the basis of time since collection; however, this procedure may substantially reduce the advantages of pooling.

#### *Selection bias*

Selection bias involves biases arising from the procedures by which the study participants are chosen from the study base. Selection bias can be avoided by including all of the study base (i.e. a cohort study) and obtaining a response rate of 100%. This

is often not practicable, but selection bias can also be controlled in the analysis by identifying factors that are related to subject selection and controlling for them as confounders. The statistical issues involved in controlling for selection bias in the analysis are essentially the same as those involved in controlling for sources of confounding (see below).

#### *Information bias*

Information bias involves misclassification of the study participants with respect to disease or exposure status. Thus, the concept of information bias refers to those people actually included in the study, whereas selection bias refers to the selection of the study participants from the study base, and confounding generally refers to non-comparability of subgroups within the study base. The various methodological issues of validity, reproducibility and stability of markers are part of the more general problem of information bias.

Non-differential information bias occurs when the likelihood of misclassification of exposure is the same for both cases and non-cases of disease (or when the likelihood of misclassification of disease is the same for exposed and non-exposed persons). Non-differential misclassification of exposure generally biases the relative risk estimate towards the null value of 1.0 (Copeland *et al.*, 1977). Hence, non-differential information bias tends to produce 'false negative' findings and is of particular concern in studies that find no association between exposure and disease.

Differential information bias occurs when the likelihood of misclassification of exposure is different between cases and non-cases (or the likelihood of misclassification of disease is different between exposed and non-exposed persons). This can bias the observed effect estimate in either direction, either towards or away from the null value.

Information bias can drastically affect the validity of a study. As a general principle, it is important to ensure that the misclassification is non-differential, by ensuring that exposure information is collected in an identical manner in cases and non-cases (or that disease information is collected in an identical manner in the exposed and non-exposed groups). In this situation, the bias is in a known direction (towards the null), and although there may be concern that not finding a significant association

(between exposure and disease) may be due to non-differential information bias, at least one can be confident that any positive findings are not due to information bias. Thus, the aim of data collection is not to collect perfect information, but to collect comparable information in a similar manner from the groups being compared, even if this means ignoring more detailed exposure information if this is not available for both groups. However, it is clearly important to collect information that is as detailed and accurate as possible, within the constraints imposed by the need to ensure that information is collected in a similar manner in the groups being compared.

In general, cross-sectional and case-control studies based on biomarkers of exposure are more prone to differential misclassification than studies based on measurement of external exposure and disease, since the biomarkers may be influenced by the disease itself. This problem is less relevant in prospective studies (or nested case-control studies) in which the marker is measured on biological material collected before the onset of disease, provided that cases diagnosed within a short interval after sample collection are excluded. The fact that the relationships between exposure, marker and disease are in most cases obscure limits the interpretation of the findings of biomarker-based studies with respect to the presence or absence of information bias.

#### Confounding

Confounding occurs when the exposed and non-exposed groups (in the study base) are not comparable due to inherent differences in background disease risk (Greenland & Robins, 1986) caused by exposure to other risk factors. The concept of confounding thus generally refers to the study base, although as noted above, confounding can also be introduced (or removed) by the manner in which study participants are selected from the study base.

If no other biases are present, three conditions are necessary for a factor to be a confounder (Rothman, 1986). First, a confounder is a factor that is predictive of disease (in the absence of the exposure under study); second, a confounder is associated with exposure in the study base; and third, a variable that is intermediate in the causal pathway between exposure and disease is not a confounder. This latter issue is of particular con-

cern in studies using biomarkers, since the identification of potential confounders depends on previous knowledge of the relationship between these and the relevant variables of exposure and outcome, and such knowledge is, for most biomarkers, very poor.

The most straightforward method of controlling confounding in the analysis involves stratifying the data into subgroups according to the levels of the confounder(s) and calculating a summary effect estimate that summarizes the information across strata. However, it is usually not possible to control simultaneously for more than two or three confounders when using stratified analysis. This problem can be mitigated to some extent by the use of mathematical modelling, but this may in turn produce problems of multicollinearity when variables which are highly correlated are entered simultaneously into the model. For example, serum levels of various micronutrients may be strongly correlated, and multicollinearity may occur when they are included in the same model. This will lead to unstable effect estimates with large standard errors, and may in fact lead to the 'wrong' micronutrient showing the strongest association (negative or positive) with disease. This may be one reason why numerous studies have shown that the consumption of green and yellow vegetables is protective against a range of cancers, but the identification of the specific dietary micronutrients involved has remained elusive (Steinmetz & Potter, 1991).

In general, control of confounding requires careful use of *a priori* knowledge, together with assessment of the extent to which the effect estimate changes when the factor is controlled in the analysis. Most epidemiologists prefer to make a decision based on the latter criterion, although it can be misleading, particularly if misclassification is occurring (Greenland & Robins, 1985). The decision to control for a presumed confounder can certainly be made with more confidence if there is supporting prior knowledge that the factor is predictive of disease.

Misclassification of a confounder leads to a loss of ability to control confounding, although control may still be useful provided that misclassification of the confounder was unbiased (Greenland, 1980). Misclassification of exposure is more problematic, since factors that influence misclassifica-

tion may appear to influence these factors may influence these factors (Greenland & Robins, 1985).

When appropriate to control confounding to assess its potential, for example, it may be possible to use a surrogate for the confounder. In this situation, it is still possible to obtain a main effect estimate controlled in the strength and direction of the effect. It may be possible to use information for a surrogate (and non-cases) in the control of confounder control.

A related approach is to use founder information (or a sample of the study). For example, in a study of biomarkers, questionnaire information from participants.

The potential for confounding in all epidemiologic studies using biomarkers. This does not reduce the validity of the study, and in some instances may actually introduce bias. For example, in a study of exposure in a group of factory workers, if are classified according to basis of industrial hygiene percentage of smoke cigarettes smoked per day in the groups with low occupational exposure are defined purely on exposure in various occupations, and these exposures to cigarette smoking will not be a confounder if workers are classified by levels, these will include information from all sources, in this case, cigarette smoke. The 'high PAH exposure' group therefore contains a higher mean (and a higher mean).

markers, since the identification depends on pre-relationship between these of exposure and out- is, for most biomarkers,

ard method of control- analysis involves stratify- according to the levels calculating a summary arizes the information usually not possible to more than two or three stratified analysis. This to some extent by the lling, but this may in multicollinearity when correlated are entered model. For example, micronutrients may be multicollinearity may d in the same model. effect estimates with ay in fact lead to the ng the strongest asso- ) with disease. This merous studies have of green and yellow st a range of cancers, pecific dietary micro- ed elusive (Steinmetz

nfounding requires edge, together with hich the effect esti- is controlled in the ts prefer to make a criterion, although it / if misclassification (Robins, 1985). The ned confounder can onfidence if there is at the factor is pre-

under leads to a loss ling, although cond- d that misclassifica- biased (Greenland, osure is more prob- ence misclassifica-

tion may appear to be confounders, but control of these factors may increase the net bias (Greenland & Robins, 1985).

When appropriate information is not available to control confounding directly, it is still desirable to assess its potential strength and direction. For example, it may be possible to obtain information on a surrogate for the confounder of interest. Even though confounder control will be imperfect in this situation, it is still possible to examine whether the main effect estimate changes when the surrogate is controlled in the analysis, and to assess the strength and direction of the change. Alternatively, it may be possible to obtain accurate confounder information for a subgroup of participants (cases and non-cases) in the study and to assess the effects of confounder control in this subgroup.

A related approach involves obtaining confounder information for a sample of the study base (or a sample of the controls in a case-control study). For example, in a study based on questionnaires, biomarkers may be used to validate questionnaire information in a subgroup of study participants.

The potential for confounding is of major concern in all epidemiological studies, including those using biomarkers. The use of biomarkers of exposure does not reduce the need to control for confounding, and in some instances the use of biomarkers may actually introduce confounding into a study. For example, in a study of lung cancer and PAH exposure in a group of factory workers, if the workers are classified according to PAH exposure on the basis of industrial hygiene monitoring, then the percentage of smokers (and the mean number of cigarettes smoked per day) will usually be similar in the groups with low, medium and high levels of occupational exposure to PAH (since these groups are defined purely on environmental levels of PAH exposure in various job categories and departments, and these exposures will usually be unrelated to cigarette smoking). In this situation, cigarette smoking will not be a major confounder. However, if workers are classified according to PAH-DNA levels, these will indicate total exposure to PAHs from all sources, including cigarette smoking. Thus, cigarette smokers will be more likely to be in the 'high PAH exposure' group, and this group will therefore contain a higher proportion of smokers (and a higher mean number of cigarettes smoked

per day) than the medium or low PAH exposure group (since some of the total PAH exposure comes from cigarette smoke). The dose-response will then be confounded by cigarette smoking, and the 'high PAH exposure' group may show a higher lung cancer risk which is not due to PAH exposure, but which is actually due to the other carcinogenic constituents of cigarette smoke (Pearce *et al.*, 1995). One solution is to stratify the analysis on cigarette smoking (as measured by questionnaire), but any confounding control is likely to be imperfect. This is even more of a problem if biomarkers are used to measure tobacco smoking because of the problems of measuring the etiologically relevant constituents of tobacco smoke (as distinct from exposure to PAHs in tobacco smoke). Once again, confounding control is likely to be imperfect and, therefore, to yield results that are still confounded and less valid than those obtained by only considering occupational exposures (using a job-exposure matrix). Furthermore, one can only control for known confounders (e.g. tobacco smoke) and cannot control for unknown confounders that may also be subject to the same types of biases described above. Therefore, it is usually preferable to avoid confounding rather than to attempt to control for it *post hoc* (which is why randomized trials are the preferred method when they are feasible). Thus, it is preferable to consider only occupational exposures to PAHs, using a job-exposure matrix, and not to attempt to measure non-occupational exposures to PAH using biomarkers. Finally, it should be stressed that the above issues have been discussed in terms of studies in which exposure is measured prospectively; the problems are even more acute when historical exposures are being assessed.

#### Random error

Random error will occur in any epidemiological study, just as it occurs in experimental studies. It is often referred to as 'chance', although it can perhaps more reasonably be regarded as 'ignorance' (Checkoway *et al.*, 1989). Even in an experimental study in which participants are randomized into 'exposed' and 'non-exposed' groups, there will be 'random' differences in background risk between the compared groups, but these will diminish in importance (i.e. the random differences will 'even out') as the study size grows. In epidemiological

studies, there is no guarantee that differences in baseline (background) risk will even out between the exposure groups as the study size grows, but it is necessary to make this assumption in order to proceed with the study (Greenland & Robins, 1986). In practice, the study size depends on the number of available participants and the available resources. Within these limitations, it is desirable to make the study as large as possible, taking into account the trade-off between including more participants and gathering more detailed information about a smaller number of participants (Greenland, 1988).

A major problem with the use of biomarkers of exposure and outcome in cancer epidemiology studies (and particularly in cohort studies) is that of small numbers. Even large multicentre cohort studies often struggle to obtain sufficient numbers to assess risks of rare cancers from occupational exposures (Fingerhut *et al.*, 1991; Saracci *et al.*, 1991). The use of biomarkers may be a major problem in this regard, since the resulting expense and complexity may drastically reduce the study size, even in community-based and nested case-control studies, and therefore greatly reduce the statistical power for detecting an association between exposure and disease. As in traditional epidemiology studies, in studies using biomarkers statistical power is a function of the prevalence of exposure and the magnitude of risk; a biomarker with low prevalence and high relative risk can be evaluated in small populations, whereas a biomarker with low prevalence and low relative risk requires a larger population. The optimal balance between precision and validity depends on a number of considerations, including the relative costs of the various exposure measurement techniques (Greenland, 1988). Thus, for an expensive biomarker to be useful it must be substantially better than less expensive and less invasive approaches. However, it has been argued that the necessary study size in some molecular epidemiology studies may be smaller than in traditional epidemiology (Hulka, 1990a,b; Hertzberg & Russek-Cohen, 1993) because of larger differences in biomarker distribution, identification of subgroups at higher risk, and the use of continuous outcome variables (Boffetta, 1995). While this is true in many cases (e.g. the detection of mutations in critical genes as a marker of increased risk of cancer), in other cases the biomarkers may show a

very low (or very high) prevalence, thus requiring large samples to detect a difference between groups (Hulka & Margolin, 1992; Rothman *et al.*, 1995). Finally, it should be noted that some biomarkers are of interest in themselves, rather than functioning as surrogates for other exposures; in particular, no alternative methods exist for measuring markers of genetic susceptibility. Nevertheless, the additional information provided by the use of such markers should still be compared with that provided by alternative, larger studies in which the marker is not used.

An additional consideration in study size estimation in studies using biomarkers is the ratio of the number of assays per individual and the number of individuals in the study (Boffetta, 1995), such as the detection of chromosomal aberrations or sister chromatid exchanges. In this case, studies must be based on adequate numbers of individuals and observations per individual (Hirsch *et al.*, 1984; Whorton, 1985). Many biomarkers show marked variation from day to day within the same individual, and the intra-individual variation may be greater than the interindividual variation (Armstrong *et al.*, 1992). It may therefore be necessary to take a large number of measurements to accurately estimate the average exposure level for each individual; otherwise it may be impossible to detect differences between individuals. For example, for 24-hour urinary sodium, the within-person variation may be three times as high as the between-person variation; it has been estimated that the misclassification resulting from taking only one sample per person would result in a true relative risk of 2.0 being reduced to an observed relative risk of 1.2 (Armstrong *et al.*, 1992). Thus, it may be necessary to take 10–15 24-hour urine samples in order to achieve reasonable accuracy in estimating average individual sodium intake levels.

#### Interaction

Interaction (effect modification) occurs when the estimate of effect of exposure depends on the level of another factor in the study base (Miettinen, 1974). The term statistical interaction denotes a similar phenomenon in the observed data. The former term will generally be used here. Interaction is distinct from confounding (or selection or information bias) in that it does not represent a bias which should be removed or controlled, but rather

a real difference in the subgroups that may be of interest. For example, in a cohort study of asthma in children, passive smoking might be a risk factor in some groups, or in males and females, or in different ethnic groups, or in males and females, or in different ethnic groups, or in different ethnic groups, or in different ethnic groups. More generally, in both groups, but the effect may vary. A typical example of interaction is in studies using biomarkers of disease due to an external agent, such as the population with a biomarker, such as the enzyme implicated in the metabolism of the agent (see Landis *et al.*, 1995). In this situation, effect estimates interpreted with consideration of statistical interaction and statistical methods used to assess factors modify either the effect, as uniformity or non-uniformity over time, or the joint effects of two or more factors. The goal of the study, this stratum-specific effect, is to proceed if statistical assessment of joint effects is not possible. Some authors (e.g. Kelsey *et al.*, 1992) argue that it is not appropriate to estimate an overall estimate of the effect, but to ignore this still, effect estimates is not valid, valid analytical methods (rate ratios) have been used in this situation (Rothman *et al.*, 1995).

Biomarkers may be useful for assessing interaction, genetic and/or environmental, genetic susceptibility, higher disease risk for than for non-susceptible (Boffetta, 1995).

However, a major problem in studies of interaction, e.g. in studies of genetic susceptibility, is that the effect of the intervention on the control study, testing:

prevalence, thus requiring a difference between groups (Rothman *et al.*, 1995). It is noted that some biomarkers are measures, rather than functions, of exposures; in particular, they exist for measuring marker activity. Nevertheless, the addition of the use of such markers compared with that proper studies in which the

variation in study size estimates of biomarkers is the ratio of the number of individuals and the number of measurements (Boffetta, 1995). In this case, studies with large numbers of individuals and few measurements per individual (Hirsch *et al.*, 1995) may show that any biomarkers show variation to day within the same individual. Individual variation may therefore be necessary for measurements to be accurate. For example, if the exposure level for a biomarker may be impossible to measure in individuals. For example, the within-person variation may be as high as the between-person variation. It has been estimated that the error from taking only one measurement may result in a true relative error of 10% to an observed relative error of 100% (Boffetta, 1992). Thus, it may be that 24-hour urine samples may be necessary for accuracy in estimating intake levels.

Interaction occurs when the effect of an exposure depends on the level of another exposure (Miettinen, 1985). Interaction denotes a departure from the expected data. The formal definition of interaction is selection or information that represents a bias that is not controlled, but rather

a real difference in the effect of exposure in various subgroups that may be of considerable interest. For example, in a cohort study of passive smoking and asthma in children, the effect estimate for passive smoking might be different in different age groups, or in males and females. The clearest example of interaction is when a factor is actually hazardous in one group and protective in another group. More generally, the risk might be elevated in both groups, but the strength of the effect may vary. A typical example of effect modification in studies using biomarkers is the estimate of the risk of disease due to an external agent in subgroups of the population with a different genetic susceptibility marker, such as the polymorphism for an enzyme implicated in the activation or detoxification of the agent (see Landi & Caporasi, this volume). In this situation, effect modification should be interpreted with considerable care, since the presence of statistical interaction may depend on the statistical methods used. In fact, all secondary risk factors modify either the rate ratio or the rate difference, as uniformity over one measure implies non-uniformity over the other. If the assessment of the joint effects of two factors is a fundamental goal of the study, this can be done by calculating stratum-specific effect estimates. It is less clear how to proceed if statistical interaction is occurring, but assessment of joint effects is not an analytical goal. Some authors (e.g. Kleinbaum *et al.*, 1982) argue that it is not appropriate in this situation to calculate an overall estimate of effect summarized across levels of the effect modifier. However, it is common to ignore this stipulation if the difference in effect estimates is not too great (Pearce, 1989). In fact, valid analytical methods (e.g. standardized rate ratios) have been specifically developed for this situation (Rothman, 1986).

Biomarkers may provide better opportunities for assessing interactions between two or more genetic and/or environmental factors. In particular, genetic susceptibility genes should produce a higher disease risk for exposed susceptible groups than for non-susceptible and non-exposed groups (Boffetta, 1995).

However, a major problem of testing for interaction, e.g. in studies involving markers of genetic susceptibility, is that it usually requires a substantial increase in study size. For example, in a case-control study, testing for interaction involves com-

paring the sizes of the odds ratios (relating exposure and disease) in different strata of the effect modifier, rather than merely testing whether the overall odds ratio is different from the null value of 1.0. The power of the test for interaction therefore depends on the numbers of cases and controls in specific strata (of the effect modifier) rather than the overall numbers of cases and controls. For example, Smith and Day (1984) give an example of a case-control study that would have to be five times larger to detect a difference between odds ratios of 1.0 and 2.0 in the two different strata of an effect modifier than it would have to be to detect an overall odds ratio of 2.0 (ignoring the effect modifier). In general, when considering possible interactions, the size of the study needs to be at least four times larger than when interaction is not considered (Smith & Day, 1984). Thus, in a study involving markers of genetic susceptibility, the gain in statistical power from considering such markers (thereby yielding higher relative risks in some strata) may be offset by the decrease in statistical power from the need to consider interactions. However, if the exposure of interest is independent from the genetic factor under study, case-case comparisons can be used to study interactions with greater statistical power (Piegorisch *et al.*, 1984).

### Conclusions

In some instances, the increasing use of biomarkers in epidemiological studies represents a major improvement in the discipline during the last years (Vineis, 1992). In many cases, biomarkers have helped to improve our knowledge of causes and mechanisms of both disease etiology and prevention. In other cases, however, it is unclear whether they represent an improvement on traditional epidemiological methods. Epidemiological studies based on biomarkers are usually more complex than traditional epidemiological studies, because information is available on a larger number of variables whose biological meaning is often poorly known. The methodological considerations involved in classical epidemiological studies on issues such as measurement of disease, measurement of exposure, selection bias, confounding, precision and interaction also apply to biomarker-based studies, and in most cases the methodological problems of this type of study do not require

solutions different from those used in classical studies. In some cases, however, the use of biomarkers may pose specific problems, which have to be addressed within the general framework of modern epidemiological methods.

#### Acknowledgements

N.P. is funded by a Professorial Research Fellowship and the Wellington Asthma Research Group is supported by a Programme Grant from the Health Research Council of New Zealand.

#### References

- Armstrong, B.K., White, E. & Saracci, R. (1992) *Principles of Exposure Measurement in Epidemiology*. New York, Oxford University Press
- Boffetta, P. (1995) Sources of bias, effect of confounding in the application of biomarkers to epidemiological studies. *Toxicol. Lett.*, 77, 235-238
- Checkoway, H., Pearce, N. & Crawford-Brown, D. (1989) *Research Methods in Occupational Epidemiology*. New York, Oxford University Press
- Copeland, K.T., Checkoway, H., McMichael, A.J. & Holbrook, R.H. (1977) Bias due to misclassification in the estimation of relative risk. *Am. J. Epidemiol.*, 105, 488-495
- Fingerhut, M.A., Halperin, W.E., Marlow, D.A., Placitelli, L.A., Honchar, P.A., Sweeney, M.H., Greife, A.L., Dill, P.A., Steenland, K. & Suruda, A.J. (1991) Cancer mortality in workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *New Engl. J. Med.*, 324, 212-218
- Forastiere, E., Agabiti, N., Dell'orco, V., Pistelli, R., Corbo, G.M., Brancato, G., Pacifici, R., Zuccaro, P. & Perucci, C.A. (1993) Questionnaire data as predictors of urinary cotinine levels among nonsmoking adolescents. *Arch. Environ. Health*, 48, 230-234
- Greenland, S. (1980) The effect of misclassification in the presence of covariates. *Am. J. Epidemiol.*, 112, 564-569
- Greenland, S. (1987) Interpretation and choice of effect measures in epidemiologic analyses. *Am. J. Epidemiol.*, 125, 761-768
- Greenland, S. (1988) Statistical uncertainty due to misclassification: implications for validation substudies. *J. Clin. Epidemiol.*, 41, 1167-1174
- Greenland, S. & Robins, J.M. (1985) Confounding and misclassification. *Am. J. Epidemiol.*, 122, 495-506
- Greenland, S. & Robins, J.M. (1986) Identifiability, exchangeability and epidemiological confounding. *Int. J. Epidemiol.*, 15, 4128
- Hertzberg, V.S. & Russek-Cohen, E. (1993) Statistical methods in molecular epidemiology. In: Schulte P. & Perera, F.P., eds, *Molecular Epidemiology: Principles and Practices*. New York, Academic Press, pp 199-216
- Hirsch, B., McGue, M. & Cervenka, J. (1984) Characterization of the distribution of sister chromatid exchange frequencies: implications for research design. *Hum. Genet.*, 65, 280-286
- Hulka, B.S. (1990a) Overview of biological markers. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, Oxford University Press, pp. 315
- Hulka, B.S. (1990b) Methodologic issues in molecular epidemiology. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, Oxford University Press, pp. 214-226
- Hulka, B.S. & Margolin, B.H. (1992) Methodological issues in epidemiologic studies using biological markers. *Am. J. Epidemiol.*, 135, 200-209
- Johnson, E.S., Parsons, W., Weinberg, C.R., Shore, D.L., Mathews, J., Patterson, D.G. & Needham, L.L. (1992) Current serum levels of 2,3,7,8-Tetrachlorodibenzo-p-dioxin in phenoxy herbicide applicators and characterization of historical levels. *J. Natl. Cancer Inst.*, 84, 1648-1653
- Kleinbaum, D.G., Kupper, L.L. & Morgenstern, H. (1982) *Epidemiologic Research. Principles and Quantitative Methods*, Belmont, CA, Lifetime Learning Publications
- McMichael, A.J. (1994) Invited commentary - 'Molecular epidemiology': new pathway or new travelling companion? *Am. J. Epidemiol.*, 140, 1-11
- McMichael, A.J. (1995) Re: invited commentary - 'Molecular epidemiology': new pathway or new travelling companion? *Am. J. Epidemiol.*, 142, 225
- Miettinen, O.S. (1974) Confounding and effect modification. *Am. J. Epidemiol.*, 100, 350-353
- Miettinen, O.S. (1976) Estimability and estimation in case-referent studies. *Am. J. Epidemiol.*, 103, 226-235
- Miettinen, O.S. (1985) *Theoretical Epidemiology*, New York, Wiley and Sons
- Muñoz, N., Bosch, F.X., Sanjose, S., Izarzugaza, I., Gili, M., Viladín, P., Navarro, C., Martos, C., Ascunce, N., Gonzalez, L.C., Kaldor, J.M., Guerrero, I., Llorca, A., Santamaría, M., Alonso de Ruiz, P., Aristizabal, N. & Shah, K. (1992) The causal link between human papillomavirus and invasive cervical cancer: a population-based case-control study in Colombia and Spain. *Int. J. Cancer*, 52, 743-749
- Pearce, N. (1992) Methodological problems of time-related variables in occupational cohort studies. *Rev. Epidemiol. Santé Publ.*, 40, S43-S54
- Pearce, N. (1993) What does the case-control study. *Int. J. Epidemiol.*
- Pearce, N.E. (1989) Analytic imlogical concepts of interaction 976-980
- Pearce, N.E., Checkoway, H. & Si lated factors as potential confounders in studies based on an occ J. Work Environ. Health, 12, 97-1
- Pearce, N., Sanjose, S., Boffetta, P. R. & Savitz, D. (1995) Limitatic posure in cancer epidemiology. :
- Piegorsch, W.W., Weinberg, C.R. : hierarchical designs for assessing tion-based case-control studies. :
- Robins, J.M., Gail, M.H. & Lubin ased selection of controls for cas hort studies. *Biometrics*, 42, 129:
- Rothman, K.J. (1986) *Modern Epi Brown*
- Rothman, N. (1995) Genetic sus studies of occupational and methodologic issues. *Toxicol. Le*
- Rothman, N., Stewart, W.E., Rai (1995) The development, valid: biomarkers for early biologic : M.L., Peeters, J.P. & Normandy, *Occupational Health: Progr Washington, D.C., Joseph Henr*
- Saracci, R. (1984) Assessing ex, the identification of disease det: Draper, M., Hemminki, K. & V: *Human Exposure to Carcinogeni* (IARC Scientific Publications national Agency for Research o
- Saracci, R., Kogevinas, M., B: Cancer mortality in an interna exposed to chlorophenoxy he nols. *Lancet*, 338, 1027-1032

1. & Cervenka, J. (1984)  
Distribution of sister chromatid  
applications for research design.

new of biological markers. In: & Griffith, J.D., eds, *Biological*  
New York, Oxford University

ologic issues in molecular  
S., Wilcosky, T.C. & Griffith,  
in *Epidemiology*, New York,  
214-226

I. (1992) Methodological issues using biological markers. 209

Weinberg, C.R., Shore, D.L.,  
T. & Needham, L.L. (1992)  
3,7,8-Tetrachlorodibenzo-*p*-  
dicarboxylic acid applicators and characteri-  
stics. *J. Natl. Cancer Inst.*, 84,

.. & Morgenstern, H. (1982)  
*Qualitative and Quantitative Methods*,  
Sage Publications

d commentary – 'Molecular  
or new travelling compan-  
-11

: invited commentary—  
w pathway or new travel-  
*miol.*, 142, 225

unding and effect modifi-  
350-353

ability and estimation in  
*Epidemiol.*, 103, 226-235

retical Epidemiology, New

S., Izarzugaza, I., Gili, M., C., Asuncion, N., Gonzalez, Lorinez, A. Santamaria, M., N. & Shah, K. (1992) The pillomavirus and invasive ased case-control study in cer, 52, 743-749

ical problems of time-re-  
nal cohort studies. *Rev.*  
\$54

Pearce, N. (1993) What does the odds ratio estimate in a case-control study. *Int. J. Epidemiol.*, 22, 1, 1189-1192

Pearce, N.E. (1989) Analytic implications of epidemiological concepts of interaction. *Int. J. Epidemiol.*, 18, 976-980

Pearce, N.E., Checkoway, H. & Shy, C.M. (1986) Time-related factors as potential confounders and effect modifiers in studies based on an occupational cohort. *Scand. J. Work Environ. Health*, 12, 97-107

Pearce, N., Sanjose, S., Boffetta, P., Kogevinas, M., Saracci, R. & Savitz, D. (1995) Limitations of biomarkers of exposure in cancer epidemiology. *Epidemiology*, 6, 190-194

Piegorsch, W.W., Weinberg, C.R. & Taylor, J.A. (1994) Non-hierarchical designs for assessing susceptibility in population-based case-control studies. *Stat. Med.*, 13, 153-162

Robins, J.M., Gail, M.H. & Lubin, J.H. (1986) More on biased selection of controls for case-control analyses of cohort studies. *Biometrics*, 42, 1293-1299

Rothman, K.J. (1986) *Modern Epidemiology*, Boston, Little, Brown

Rothman, N. (1995) Genetic susceptibility biomarkers in studies of occupational and environmental cancer: methodologic issues. *Toxicol. Lett.*, 77, 221-225

Rothman, N., Stewart, W.F., Rabkin, C.S. & Schulte, P.A. (1995) The development, validation, and application of biomarkers for early biologic effects. In: Mendelsohn, M.L., Peeters, J.P. & Normandy, M.J., eds, *Biomarkers and Occupational Health: Progress and Perspectives*, Washington, D.C.: Joseph Henry Press, pp. 109-115

Saracci, R. (1984) Assessing exposure of individuals in the identification of disease determinants. In: Berlin, A., Draper, M., Hemminki, K. & Vainio, H., eds, *Monitoring Human Exposure to Carcinogenic and Mutagenic Agents*, (IARC Scientific Publications No. 59), Lyon, International Agency for Research on Cancer

Saracci, R., Kogevinas, M., Bertazzi, P., *et al.* (1991) Cancer mortality in an international cohort of workers exposed to chlorophenoxy herbicides and chlorophenols. *Lancet*, 338, 1027-1032

Schatzkin, A., Freedman, L. & Schiffman, M. (1993) An epidemiologic perspective on biomarkers. *J. Int. Med.*, 233, 75-79

Schwartz, G.G. (1990) Chromosome aberrations. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, Oxford University Press. pp. 146-172

Siemiatycki, J., Wacholder, S., Dewar, R., Wald, L., Bégin, D., Richardson, L., Rosenman, K. & Gérin, M. (1988) Smoking and degree of occupational exposure: are internal analyses in cohort studies likely to be confounded by smoking status? *Am. J. Ind. Med.*, 13, 59-69

Smith, P.G. & Day, N.E. (1984) The design of case control studies: the influence of confounding and interaction effects. *Int. J. Epidemiol.*, 13, 356-365

Steinmetz, K.A. & Potter, J.D. (1991) Vegetables, fruit and cancer: I. epidemiology. *Cancer Causes Control*, 2, 325-357

Vineis, P. (1992) Uses of biochemical and biological markers in occupational epidemiology. *Rev. Epidem. Sante Publ.*, 40, S63-S69

Whorton, E.B. (1985) Some experimental design and analysis considerations for cytogenetics studies. *Environ Mutagen*; 4 (Suppl), 9-15

Wilcosky, T.C. & Griffith, J.D. (1990) Applications of biological markers. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, Oxford University Press, pp. 16-27

Willett, W. (1990) *Nutritional Epidemiology*. New York, Oxford University Press

*Corresponding author*

N. Pearce

Wellington Asthma Research Group,  
Wellington School of Medicine, PO Box 7343,  
Wellington, New Zealand